

BIROn - Birkbeck Institutional Research Online

Dehaene-Lambertz, G. and Gliga, Teodora (2004) Common neural basis for phoneme processing in infants and adults. *Journal of Cognitive Neuroscience* 16 (8), pp. 1375-1387. ISSN 0898-929X.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/31606/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

Common Neural Basis for Phoneme Processing in Infants and Adults

G. Dehaene-Lambertz^{1,2} and T. Gliga¹

Abstract

■ Investigating the degree of similarity between infants' and adults' representation of speech is critical to our understanding of infants' ability to acquire language. Phoneme perception plays a crucial role in language processing, and numerous behavioral studies have demonstrated similar capacities in infants and adults, but are these subserved by the same neural substrates or networks? In this article, we review event-related potential (ERP) results obtained in infants during phoneme

discrimination tasks and compare them to results from the adult literature. The striking similarities observed both in behavior and ERPs between initial and mature stages suggest a continuity in processing and neural structure. We argue that infants have access at the beginning of life to phonemic representations, which are modified without training or implicit instruction, but by the statistical distributions of speech input in order to converge to the native phonemic categories. ■

INTRODUCTION

Cognitive capacities, such as language, mathematics, or music are highly developed in humans as compared to animals. Numerous studies have found precursors of these capacities in infants: For example, infants are able to discriminate sentences in different languages (Mehler et al., 1988), distinguish sets of objects based on their numerosity (Feigenson, Carey, & Spelke, 2002), or recognize known faces (Bushnell, 1982). These abilities are not so different from those of other animals. Tamarin monkeys are also able to discriminate two human languages (Ramus, Hauser, Miller, Morris, & Mehler, 2000) and two quantities of items (Hauser, Dehaene, Dehaene-Lambertz, & Patalano, 2002), and rhesus monkeys respond to particular faces (Parr, Winslow, Hopkins, & de Waal, 2000). In a few years, however, children surpass these animals. The challenge that developmental psychology faces is to explain how the infant brain becomes able to assume complex cognitive functions and why these functions rely on specialized neuronal networks that are reproducibly located across human adults. Because the human genome cannot code directly for human complex cognitive capacities, exposure to the human environment may take advantage of preexisting biases in human brain functional architecture. Therefore, we need first to determine what the essential properties of a given processing system are in adults and how similar are the capacities present in infants. Second, we have to examine whether similar capacities are subserved by analogous neural organization and to understand how initial biases in brain organization could be shaped by the human environment to give rise to the mature state.

In this article, we will consider the case of phoneme perception.¹ By studying the cerebral bases of phoneme perception at the initial and mature stages, we can examine whether there is continuity or discontinuity between infants and adults and whether the linguistic environment shapes an existing network or creates a new functional organization.

Phoneme Perception in Adults

One main characteristic of phoneme perception is the capacity to identify the same phoneme across different acoustical realizations. These differences are related to speakers' vocal tract, speech rate, speech mode, or to the surrounding phonemes. Although many acoustical variations are not phonetically distinctive, sharp transitions (which are very robust across subjects) change the phonemic category. Normalization and categorical perception are two important properties essential for speech comprehension. Duplex perception and integration of visual information leading to MacGurk² or ventriloquy effects are also properties associated with phoneme perception. None of these characteristics is strictly specific to phoneme perception and most can be observed in the perception of other sounds. For example, the perception of tones differing on the duration of their initial frequency transition is categorical (Pisoni, 1977), and duplex perception is described for nonspeech sounds (Fowler & Rosenblum, 1990). However, no other type of auditory processing relies so heavily and systematically on these properties than speech processing.

Another important characteristic of phoneme perception in adults is that it depends on the subjects' native language. Adults have difficulties in discrimi-

¹INSERM-CEA, ²Centre Hospitalier Universitaire Bicêtre

nating foreign phonemic contrasts that are not present in their native language, such as /r/-/l/ for Japanese speakers (Goto, 1971), /e/-/ɛ/ for Spanish speakers (Pallier, Bosch, & Sebastian, 1997), retroflex versus dental /d/ for English speakers (Werker & Tees, 1984b), among many examples. The influence of native language on speech perception is not limited to the phoneme repertoire of the language, but concerns all aspects of phonology. Acoustic cues coding suprasegmental information, such as tones or stress, and phonotactics, the set of rules that govern how phonemes are combined within words, are processed differently by subjects from different languages. For example, Japanese adults have difficulties in discriminating two pseudowords like /ebzo/ and /ebuzo/. They perceive an illegal word like /ebzo/ as /ebuzo/, inserting an illusory vowel between consonants because Japanese does not allow for consonant clusters (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999).

Phoneme Perception in Infants

Behavioral studies have demonstrated striking similarities between the phoneme perception capacities of infants and adults. Like adults, infants perceive phonemes categorically along acoustic dimensions such as voice onset time (VOT) (Eimas, Siqueland, Jusczyk, & Vigorito, 1971), place of articulation (Werker, Gilbert, Humphrey, & Tees, 1981), and others. They are able to normalize across irrelevant acoustic variations such as those related to different voices or pitches (Jusczyk, Pisoni, & Mullennix, 1992; Kuhl, 1983), or coarticulation context (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy, & Mehler, 1988). They display duplex perception (Eimas & Miller, 1992). They are able to detect a match between a vowel sound and the correct corresponding face movements (Patterson & Werker, 2003; Kuhl & Meltzoff, 1984), and are sensitive to the MacGurk effect (Rosenblum, Schmuckler, & Johnson, 1997). The linguistic environment rapidly modifies these initial capacities. Around 6 months, infants show greater sensitivity to vowel categories of their native language (Polka & Werker, 1994; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992) and around 10 months, to consonants (Kuhl, Tsao, & Liu, 2003; Werker & Tees, 1984a). Although no experiment has directly demonstrated a decrease of discrimination capacities for other aspects of phonology than those related to the repertoire of phonemes, several experiments have shown that around 9–10 months of age, infants use their knowledge of phonotactics to segment words (Mattys & Jusczyk, 2001) or to distinguish list of words differing on the probability of sequences of phonemes that constitute them (Jusczyk, Luce, & Charles-Luce, 1994). Thus, at the end of the first year of life, infants certainly have access to phonological representations similar to those present in adults. Are the same behavioral capacities in infants and adults sustained by the same brain organization? What is the initial state of or-

ganization of the human brain and how can it facilitate language acquisition?

METHODOLOGICAL ISSUES: MISMATCH RESPONSES AS A TOOL TO EXPLORE AUDITORY REPRESENTATIONS

Event-related potentials (ERPs) are easily recorded even in young children and elicitation of mismatch responses in oddball paradigms is a powerful tool to access the properties of the representations computed by the cortex from a sound. It has been observed in single-cell recordings in the visual (Miller, Li, & Desimone, 1991) and auditory cortex (Ulanovsky, Las, & Nelken, 2003) that the response of neurons decreases with repetitive exposure to the same stimulus. When a new or deviant stimulus is presented, it elicits a response of a new set of neurons. Using this property called repetition suppression, it is possible to isolate cells that code different properties of the stimulus, such as object identity, independently of its size or location (Miller et al., 1991). Of course, the spatial resolution of brain imaging techniques is too low to distinguish activity at the scale of single neurons or columns of neurons that are the coding elements of the nervous system. However, it is possible to access this code at a macroscopic level by contrasting the decreased activity recorded when the stimulus is repeated with the renewed activity due to the response of even a close set of neurons in the same brain volume (see Naccache & Dehaene, 2001, for a complete discussion of the priming method in brain imagery). This difference in activity is certainly the source of mismatch responses recorded in ERP oddball paradigms. By manipulating what elicits a mismatch response, that is, what counts or not as a repetition for a given neural network, it is possible to infer the format of representation that is computed by the network. For example, a network computing a phonemic representation should habituate to the repetition of the same phoneme independently of the speaker and recover when the phoneme is changed, whereas a network coding for speaker should be affected by a voice change but not by a phoneme change. Although mismatch negativities (MMNs) observed in auditory oddball paradigms have been interpreted as a specific auditory phenomenon related to a trace of the repeated stimulus in echoic memory (Näätänen, 2001), these responses are better understood in this more general context of repetition suppression. The experimental conditions, by selecting the level of representation that is targeted by the repetition context, generate different MMNs recorded with different topographies and latencies because of the activation of different networks within the temporal lobes. Some MMNs are thus related to the computation of simple acoustic cues (e.g., duration, intensity of the stimulus: Giard et al., 1995), but others imply that multiple cues, such as conjunction of features

(Takegata, Paavilainen, Näätänen, & Winkler, 1999), arbitrary rules (Horvath, Czigler, Sussman, & Winkler, 2001), lexical and grammatical status (Shtyrov & Pulvermuller, 2002; Pulvermuller et al., 2001) have been integrated into more complex representations. Beyond the identification of the code computed by the network, which is obtained by manipulating the properties of the repeated stimulus, crucial information can be obtained about the speed of the computation, the brain areas involved and its temporal duration by measuring the latency and the topography of the mismatch responses, as well as the maximal temporal spacing between stimuli that still permits the recording of a mismatch response.

If similar representations are computed in infants and adults, mismatch responses in both populations should present the same functional properties. However, we should not expect to record the mismatch response at the same latency and with the same topography in infants than in adults. As can be seen in Figure 1, the brain response changes rapidly in morphology, topography, and latency during the first months of life (Kushnerenko, Ceponiene, Balan, Fellman, Huotilaine, et al., 2002; Novak, Kurtzberg, Kreuzer, & Vaughan, 1989; Barnett, Ohlrich, Weiss, & Shanks, 1975, for a description of auditory ERP maturation) because of changes within auditory cortices, such as myelinization, differential maturation of the cortical layers, folding of the cortical surface, and modifications of intracortical connections, but also to more general changes that may affect electrical transmission such as differential expansion of brain areas, closure of the fontanel, ossification of the skull, and others. For example, N1 in adults is described as a broad negativity recorded 100 msec after a sound at the vertex and surrounding electrodes. It has been related to the activation of numerous cerebral generators mainly in the planum temporale (Eggermont & Ponton, 2002). If such a component is not recorded during the first months of life, it does not mean that sounds cannot access an immature auditory cortex. Functional MRI studies found activations to sound in cortical regions that are grossly similar in infants and in adults (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002; Altman & Bernal, 2001; Anderson et al., 2001). Therefore, descriptive properties of ERP components, such as latencies and topographies determined from a few electrodes, are often not relevant to compare the cortical areas activated and their computational properties at different ages. We should rather rely on an experimental design that targets a precise computation on the stimulus and on dipole modeling of activated brain areas using high-density coverage of the scalp.

Cerebral Bases of Phoneme Perception in Adults

What are the cerebral bases of phoneme perception in adults? Among the representations of sound that may be

computed and are accessible to mismatch paradigms, can we identify a phonemic representation? A network coding such a representation should display the same properties as those described in behavioral paradigms for phoneme perception, for example, categorical perception. Indeed, a change of consonant occurring within a phonemic category elicits a significantly smaller MMN, or no MMN, than does a similar change that crosses a phonemic boundary (Sharma & Dorman, 1999; Dehaene-Lambertz, 1997). The MMN is also sensitive to the subjects' native language. The same change in a series of syllables induces an MMN in subjects for whom the change is linguistically pertinent in their language, but not (or a significantly weaker one) in subjects for whom the change is not linguistically pertinent (Sharma & Dorman, 2000; Winkler et al., 1999; Dehaene-Lambertz, 1997; Näätänen et al., 1997). This effect of native language is not limited to the repertoire of phonemes but encompasses the entire phonology of the language. For example, phonotactics influence MMN responses: A change in pseudowords (e.g., ebuzo to ebzo) induces an MMN in French subjects but not in Japanese subjects (Dehaene-Lambertz, Dupoux, & Gout, 2000). Finally, the visual integration of incongruent articulatory movements during phoneme perception induces a mismatch response even when the auditory stimulus itself does not change (Colin et al., 2002; Sams et al., 1991). These results demonstrate that oddball paradigms are able to reveal a level of representation that displays properties representative of phoneme perception, namely, categorical perception, speaker normalization (Dehaene-Lambertz et al., 2000), influence of the native language, and visuoauditory integration.

By virtue of the temporal resolution provided by ERPs, we can investigate whether acoustic and linguistic representations of a syllable are computed sequentially or in parallel. When an MMN is present for an acoustic change in speech stimuli (e.g., a within-category change), it never precedes a phonemic MMN (Phillips et al., 2000; Rivera-Gaxiola, Csibra, Johnson, & Karmiloff-Smith, 2000; Sharma, Marsh, & Dorman, 2000; Winkler et al., 1999; Dehaene-Lambertz, 1997; Näätänen et al., 1997). These results confirm Whalen and Liberman's (1987) hypothesis that the representations computed from speech are immediately phonemic with no intermediate acoustic representation. It is important to keep in mind that syllables are not only linguistic, but also acoustic stimuli. The very same dimensions, for example, VOT or formant transition shape, are integrated by a phonemic network into a phonemic representation, but can also be computed by acoustic networks to create acoustic representations along these dimensions. The joint representation of speech within both an acoustic representation and a phonemic representation is exemplified by the perception of sine wave stimuli that can be heard both linguistically and nonlinguistically (Liebenthal, Binder, Piorkowski, & Remez, 2003; Remez, Pardo,

Piorkowski, & Rubin, 2001). Acoustic representations are also used to discriminate within-category changes and foreign contrasts. However, these acoustic networks are less efficient and compute less stable memory traces than the phonemic network as demonstrated by our poor performance in perceiving these changes. It is also possible that the phonemic network, once activated by speech, exerts an inhibitory influence over these auditory representations to prevent interference from nonlinguistically pertinent differences (Liebenthal et al., 2003; Liberman, Isenberg, & Rakerd, 1981; Dehaene-Lambertz, Pallier, Serniclaes, Sprenger-Charolle, & Dehaene, submitted).

Where is the phonemic network located? Dipole models of the phonemic MMN suggest that this response originates from the planum temporale and is asymmetric, favoring the left hemisphere (Tervaniemi et al., 1999; Näätänen et al., 1997). This is congruent with neuropsychological observations that have established that correct phoneme perception depends on the integrity of the posterior part of the left perisylvian regions (Dronkers, Redfern, & Knight, 2000; Caplan, Gow, & Makris, 1995). Brain imaging studies (Jacquemot, Pallier, LeBihan, Dehaene, & Dupoux, 2003; Vouloumanos, Kiehl, Werker, & Liddle, 2001; Binder et al., 2000) and discrimination deficits elicited by electrical stimulation in patients with implanted subdural electrodes arrays (Boatman, Lesser, & Gordon, 1995) have confirmed the crucial role of the posterior and superior part of the left temporal lobe and of the left inferior parietal gyrus in phoneme perception in adults. The asymmetry favoring the left hemisphere for phoneme perception has been related by some authors to the fact that the left hemisphere is specialized to process stimuli with fast temporal changes, which are frequent in speech (Zatorre, Belin, & Penhune, 2002). Other authors suggest a genetic predisposition of the left hemisphere to process all aspects of language, from phoneme perception to syntax. fMRI studies have shown that activation in some left regions does not depend only on the physical characteristics of the stimuli but rather on their linguistic value. Significant differences are observed between subjects with different native languages in the left temporal and parietal lobe (Jacquemot et al., 2003; Klein, Zatorre, Milner, & Zhao, 2001) or in the left frontal region (Gandour et al., 2002). Using fMRI with sine wave analogues of /ba/ /da/, Dehaene-Lambertz et al. (submitted) observed that brain activations are always significantly asymmetric favoring the left side, even when sine wave stimuli were perceived as whistles, confirming that the left hemisphere is better suited to process stimuli with fast transitions. However, two brain areas in the left hemisphere, the posterior part of the superior temporal sulcus and the supramarginal gyrus, were specifically activated when the sine waves were perceived as syllables. These results show that the acoustical properties of the signal are not sufficient to explain brain activations to speech stimuli. For equal stimulus characteristics,

some networks in the left hemisphere are specifically correlated with the perception of stimuli as speech.

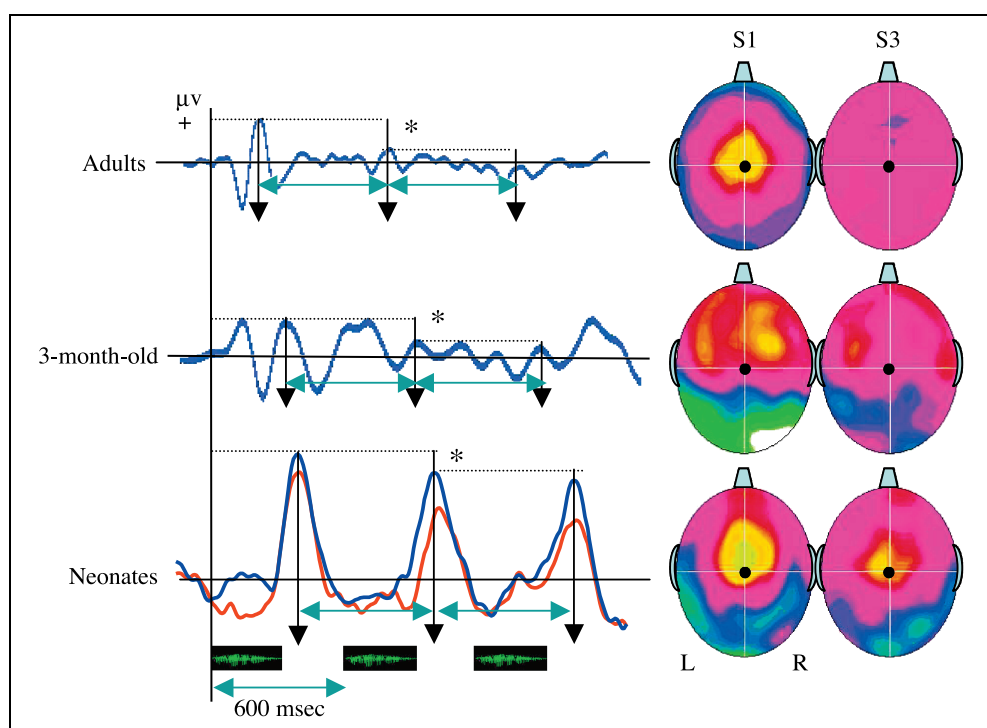
To summarize, ERP experiments in adults suggest that several representations are computed in parallel, generating different mismatch responses around 100–300 msec. One of these corresponds to a phonemic representation that presents the same functional properties as those identified in behavioral studies of speech processing. Neuropsychology and brain imaging studies locate this network in the left temporoparietal junction (Wernicke's area). Can we isolate a network in infants demonstrating the same properties as the one present in adults?

Cerebral Bases of Phoneme Perception in Infants

We have hypothesized that mismatch responses are related to a decrease in activity due to repetition. In infants, as in adults (Woods & Elmasian, 1986), repetition of the same stimulus induces a decrease in amplitude of the event-related response (Dehaene-Lambertz & Dehaene, 1994), and this decrease in amplitude is larger when the interstimulus interval (ISI) is shorter (Leppanen, Pihko, Eklund, & Lyytinen, 1999). The effect of repetition is similar across ages, inducing a sharp and significant decrease between the first and second presentation (Figure 1). When syllables belonging to the same phonemic category are produced by different speakers, the same decrease in amplitude is observed, suggesting that a normalization process is involved (Dehaene-Lambertz, Pena, Christophe, & Landrieu, 2004; Dehaene-Lambertz & Pena, 2001) and that a phonemic representation is accessed (Figure 1).

The introduction of a new stimulus induces a mismatch response that displays functional similarities with adults' MMN. It is recorded even when attention is not directed toward the stimuli, either because infants are looking at interesting visual stimuli to keep them quiet during recording (Dehaene-Lambertz & Dehaene, 1994) or because they are asleep (Dehaene-Lambertz & Pena, 2001; Alho, Saino, Sajaniemi, Reinikainen, & Näätänen, 1990) or even in a coma vigil (Dehaene-Lambertz et al., 2004). When babies are awake, this response precedes a late frontal negativity, beginning around 800 msec (Friederici, Friedrich, & Weber, 2002; Dehaene-Lambertz & Dehaene, 1994; Kurtzberg, Stone, & Vaughan, 1986). Because this late negativity is recorded after unexpected visual and auditory stimuli (Nelson & deRegnier, 1992; Kurtzberg et al., 1986; Courchesne, 1983) and is not present when babies are asleep (Friederici et al., 2002), it is seen as reflecting general attention-orienting processes. In adults, the auditory mismatch response is followed by a P300 when subjects have to detect the change of stimulus. Although no study has recorded behavioral responses and ERPs in the same session in infants, this two-step progression, auditory mismatch

Figure 1. Repetition of the same syllable induces a decrease in ERPs in adults (top), 3-month-old infants (middle), and neonates (bottom), already significant between the first and the second syllable (*). Left: grand-average across subjects of recordings at the vertex. For neonates, two waveforms are presented showing similar habituation when the syllable is physically identical (blue line) and when the syllable is produced by different speakers (red line). Right: topographic maps of the voltage at the same time point (arrow on the waveforms) for the first (S1) and the third syllable (S3). Although morphology and latencies of the ERPs are very different across ages, the same repetition priming effect is observed (adapted from Dehaene-Lambertz, 1997; Dehaene-Lambertz & Baillet, 1998; Dehaene-Lambertz & Pena, 2001).



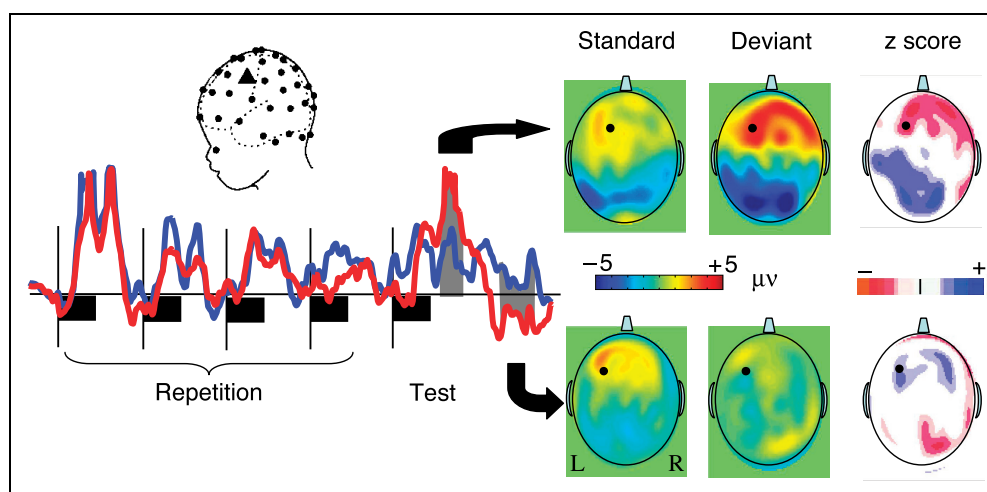
response followed by an amodal attentional component, is reminiscent of the two stages observed in adults after a deviant attended stimulus (Figure 2).

As in adults, there is not one, but several mismatch responses depending on the stimulus and on the feature of the stimulus that changes. For example, within the same infants, the response to a change in voice has a different topography than the response to a change in phoneme, demonstrating that voice and phoneme category, among other features of the stimulus, have been coded in parallel by different networks (Dehaene-Lambertz, 2000). The mismatch response to syllables displays properties essential for phoneme perception such

as normalization across speakers. In neonates, a similar mismatch response is found when the same physical stimulus is repeated before the phonemic change and when each syllable is produced by a different speaker and thus is physically different (Dehaene-Lambertz & Pena, 2001). This implies that the representation computed from the speech signal in this case is independent of the acoustic variations related to changes of intensity, duration, prosodic contour and timbre of voice, and that it is only affected by a change of phoneme.

Categorical perception is demonstrated by the fact that when syllables varying along a place of articulation (/ba/ to /da/) are used, the mismatch response is larger

Figure 2. Grand average recorded from a left frontal electrode (● on the maps) in 16 three-month-old infants during an entire trial for the standard (blue line) and deviant condition (red line). A first mismatch response originating from the temporal lobes is recorded around 400 msec. A second slow wave develops after 600 msec over the frontal areas (adapted from Dehaene-Lambertz & Dehaene, 1994).



for a change that crosses the phonemic boundary (AC) than for a change of similar amplitude within the phonemic category (WC) (Dehaene-Lambertz & Baillet, 1998). As in adults, the mismatch response to the within-category change does not precede the mismatch response to across-category change, but occurs at the same latency. Furthermore, their topography differs, suggesting that different brain areas are involved (Figure 3). The mismatch response to the within-category change is more central than the response to the across-category change, which is present above the right frontal and left occipital areas. When a dipole model of these responses is computed, the proposed sources are located more posterior and dorsal for the linguistic response (AC) than for the acoustical response (WC) (Dehaene-Lambertz & Baillet, 1998). In similar mismatch paradigms using fMRI, adults display greater activations during phonemic discrimination than during acoustic discrimination in the posterior part of the superior temporal sulcus and in the adjacent parietal region (supramarginal and angular gyrus) (Jacquemot et al., 2003; Dehaene-Lambertz et al., submitted). The more important contribution of posterior perisylvian areas in phonemic coding is thus compatible with the shifting of the source localization found in infants (Figure 3). This result suggests that as in adults, distinct acoustic and phonemic representations are computed in parallel from the same syllable.

Dipole modeling of the mismatch response to syllables locates the brain areas activated in infants in the posterior part of the temporal lobe (Dehaene-Lambertz & Baillet, 1998; Dehaene-Lambertz & Dehaene, 1994). In adults, this response is asymmetric, originating mainly from the left hemisphere. In infants, both ERPs to CV syllables and mismatch responses are asymmetric, favoring the left hemisphere, but a similar asymmetry is observed for acoustic processing. For example, the difference in the timbre of two tones elicits an asymmetric mismatch response, which is larger above the left hemisphere (Dehaene-Lambertz, 2000; Duclaux, Chalmel, Collet, Rouillet-Solignac, & Revol, 1991). In a dipole model, the vector is longer on the left side than on the right side, both for a within-category and an across category change (Figure 3). Thus, the left auditory cortex seems to present a higher responsiveness than the right for auditory stimuli in general and not specifically for linguistic stimuli. A similar pattern is observed with fMRI in 3-month-old infants (Dehaene-Lambertz et al., 2002). Forward, but also backward speech, activates the left temporal areas more than the right, and no interaction was observed between the linguistic pertinence of the stimuli and the hemisphere activated. However, both stimuli contain high-frequency transitions. Using optical topography, Sato et al. (2003) compared the lateralization of the bold response in auditory areas to a pitch change (/itta/ vs. /itta?/) and to a phoneme change (/itta/ vs. /itte/) during the first year of life. A

difference in laterality between the two conditions was not found until the age of 11–12 months on. More studies with high spatial resolution such as fMRI or optical topography are needed to confirm whether the left hemisphere is more responsive than the right to any sound, or only to sounds with fast transitions. The higher responsiveness of the left hemisphere to sounds during the first months of life while babies are intensively learning their native language might be one developmental bias that pressures language processing toward the left hemisphere. This leftward asymmetry described in normal infants appears to be more a bias than to be related to incapacity of the right hemisphere to process syllables. Patient LG suffers from a left sylvian infarct that occurred at birth. Tested 3 weeks after the acute lesion, her right hemisphere was able to discriminate a change of phoneme even when several speakers produced the stimuli (Dehaene-Lambertz et al., 2004). On the other hand, patient SD, with an analogous right hemisphere lesion, was easily able to discriminate a change in timbre with his left hemisphere. In adults, such a lesion would have induced a severe deficit in this timbre discrimination task based on spectral computation (Figure 4). These results suggest a different lateralization pattern in infants and adults.

The influence of native language on phoneme perception is not noticeable in behavior before 5–6 months of age. A phonemic mismatch response should be affected in a similar way. Only one published experiment has explored the perception of foreign phonemic contrasts with ERPs. Cheour, Ceponiene, et al. (1998) recorded the response to a change of vowel from /e/ to /ö/ and /o/ in Estonian and Finnish infants. /o/ is not present in Finnish. Finnish adults, contrary to Estonians, show a smaller MMN for /o/ than for /ö/ although the acoustic distance is wider for the change /e/–/o/ than for /e/–/ö/. At 1 year of age, the mismatch response recorded for the foreign vowel /o/ is smaller in Finns than in Estonians, whereas at 6 months, the response is similar in both populations. In adults, the duration of the mismatch response is strongly correlated with discrimination performance. It is thus surprising to record a similar mismatch response at 6 months in Finns and Estonians, especially considering that numerous behavioral experiments have shown that vowel perception is already affected by native language at this age (Polka & Werker, 1994; Kuhl et al., 1992). However, we do not have behavioral data on the precise contrast used in Cheour et al.'s electrophysiological experiment. The typical development often presented in the literature (effects of native language at 6 months for vowel perception and at 10 months for consonant perception) may be oversimplified. Some contrasts may be obscured at different ages depending on the adjacent phonemes in the phonemic space of the native language or on the robustness of the phonemic boundary (Maye, Werker, & Gerken, 2002; Best, McRoberts, & Sithole, 1988).

Figure 3. Categorical perception in 3-month-old infants along a synthetic continuum (/ba/ to /da/). Control (/ba1/ /ba1/ /ba1/ /ba1/), within-category change (/ba2/ /ba2/ /ba2/ /ba1/), and across-category change (/da/ /da/ /da/ /ba1/) trials were randomly presented. The physical distance was similar along the continuum for the within-category change and the across-category change (top). Voltage cartographies of the same test syllable (/ba1/) at the maximum of the mismatch response (middle) and dipole models of the active brain regions (bottom) for the three conditions. The dipole for the across-category change is more posterior and dorsal than for the within-category change (adapted from Dehaene-Lambertz & Baillet, 1998).

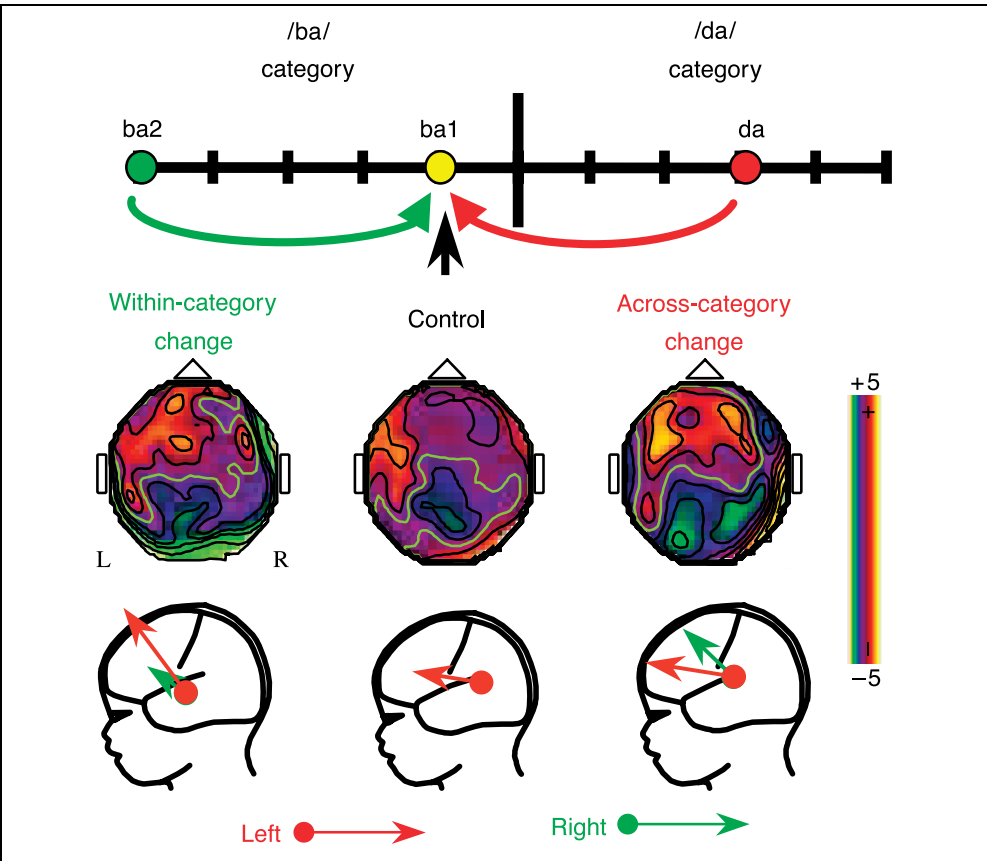
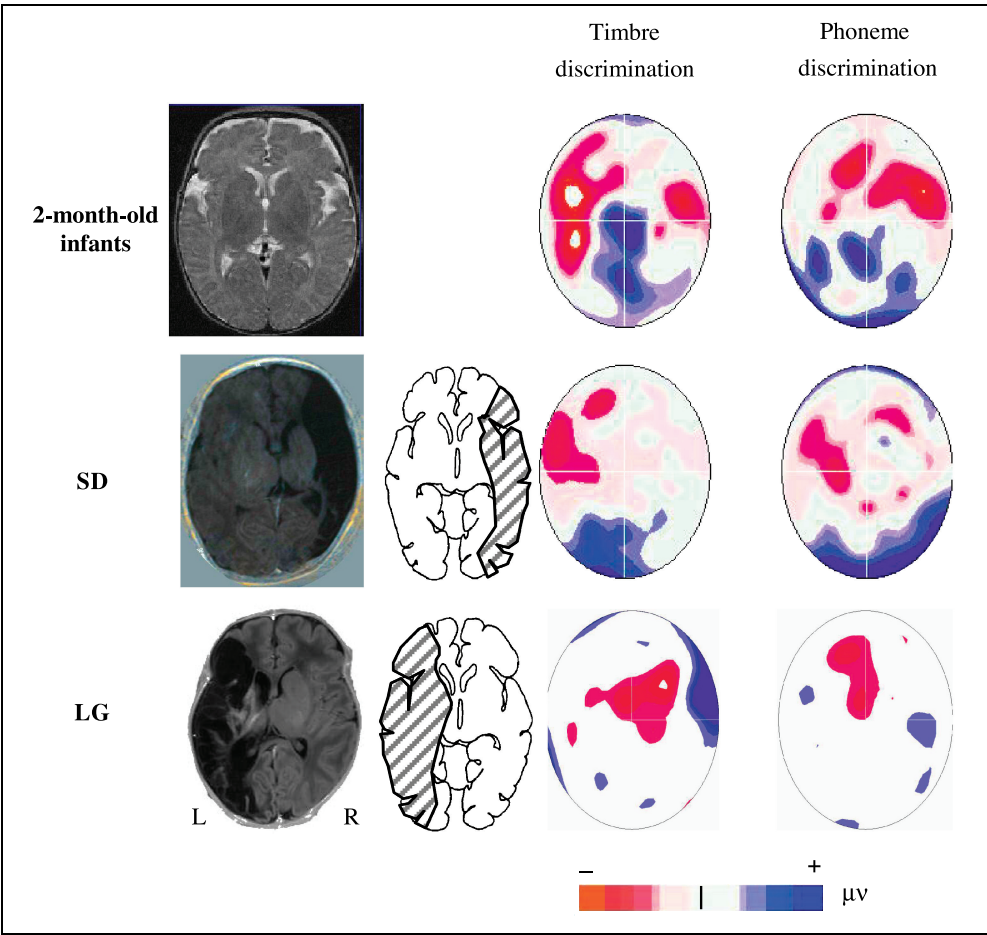


Figure 4. Mismatch responses for a phoneme discrimination task and a timbre discrimination task in normal 2-month-old infants and in two patients, SD and LG. These two babies had suffered from an important hemispheric lesion due to a sylvian infarct at birth, on the left side for LG and on the right side for SD. Despite their lesions, they show a mismatch response in the contralateral healthy hemisphere (adapted from Dehaene-Lambertz et al., 2004; Dehaene-Lambertz & Pena, 2001, for the normal infants, Dehaene-Lambertz, 1997, for LG, and unpublished data for SD).



Second, it is possible that the mismatch response was in fact already differentiable in both populations at 6 months, but that this was not observed because the spatial sampling of the voltage on the scalp was scarce (only Cz and Fz were recorded) and because the acoustical distance of the change in this experiment was confounded with its linguistic pertinence. Further experiments testing other phonemic contrasts and combining behavioral and ERP results should be carried out to clarify these points. However, this result shows at least that the difference in behavior observed across subjects from different linguistic backgrounds from the end of the first year of life on is based on a difference in the early representations computed from the speech signal.

Differences Between Infants' and Adults' Mismatch Responses

Using the same ERP paradigm in infants as in adults, we have observed a mismatch response to a change of syllable that displays similar functional properties in both populations: categorical perception and normalization. It is elicited automatically even when infants are asleep, and this response precedes an amodal frontal response when infants are awake, reminiscent of what is observed in adults processing an attended or unattended stimulus. Dipole modeling suggests a temporal origin as in adults. Although the functional properties of mismatch responses and their corresponding active brain areas appear to be similar in infants and adults, suggesting similar phonemic representations, it should be noted that the latency of the first mismatch effect is delayed in infants relative to adults. For instance, the maximum of the mismatch response for the same stimuli /ba/ /da/ is at 280 msec in adults versus 454 msec in 3-month-old babies (Dehaene-Lambertz, 1997; Dehaene-Lambertz & Baillet, 1998). Secondly, in most of the experiments cited above, the polarity of the mismatch response is reversed. Whereas in adults the mismatch response is described as a frontal negativity (from which the name mismatch negativity arises) with a positivity over the temporal regions, in our experiment with infants the mismatch response is positive over frontal areas and negative over temporo-occipital areas along a similar right frontal–left posterior axis as in adults (Dehaene-Lambertz, 1997; Dehaene-Lambertz & Baillet, 1998). Several other authors have also recorded a similar mismatch frontal positivity in infants (Winkler et al., 2003; Friederici et al., 2002; Morr, Shafer, Kreuzer, & Kurtzberg, 2002; Pihko et al., 1999; Duclaux et al., 1991; Alho, Sajaniemi, Niittyvuopio, Saino, & Näätänen, 1990) and in toddlers (Maurer, Bucher, Brem, & Brandeis, 2003). We have already discussed the hypothesis that activity of the same networks might result in different voltage topographies on the scalp across ages because of the structural changes occurring in the brain

during development. However, some authors have observed frontal negativities (Cheour, Alho, et al., 1998; Cheour, Ceponiene, et al., 1998; Cheour et al., 2002) and have suggested a strict similarity of the mismatch response across ages (Cheour, Alho et al., 1998). These negativities, however, are usually very long, several hundred milliseconds, and do not show the same sharp temporal profile as in adults (e.g., compare the 100-msec duration of the MMN for a vowel contrast in adults and the 400-msec duration for the same contrast in 12-month-old infants (Cheour, Ceponiene et al., 1998; Näätänen et al., 1997). Moreover, shortcomings that do not facilitate the interpretation of the results across ages are sometimes present in infant studies. First, the reference electrode affects the visualization of ERPs, and it should be chosen to be as neutral as possible, either by using a distant electrode that is not affected by the activity of the voltage generators or by computing an average reference if a large number of electrodes is used. In the latter case, subtracting the average voltage from the voltage recorded at each electrode creates reference-free voltages (Bertrand, Perrin, & Pernier, 1985). However, EEG recordings in infants have often used a mastoid reference. As can be seen in Figure 1, the auditory average reference ERP is negative above the mastoid in infants, and the mastoid should not be considered as neutral. Thus, a comparison across ages should be made cautiously when the reference is located on an active site. Second, in order to measure a mismatch response in classical oddball designs, the response to a deviant presented 20% of the time is compared with the same stimulus presented in another block 100% of the time. The block effect is thus confounded with time. Time could induce habituation of low-level processing or changes in listening strategy even during passive tasks. This could be worse in children, who grow increasingly restless as time passes, with possible increases in movement artifacts, changes in attention, and others. Often, this second block is not recorded in infants, and the comparison is thus done between different standard and deviant stimuli within the first block. In this case, if a difference is recorded, it could be related either to a genuine discrimination process or to differences in ERP morphologies and latencies due to the physical differences between the two stimuli. For example, the latency and the morphology of N1 are affected by VOT. A comparison between two stimuli of a VOT continuum would show differences, but these are not systematically associated with the perception of a difference (Sharma et al., 2000). The morphology of infants' ERPs is also affected by the physical characteristics of the stimulus, such as its duration (Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002), its intensity envelope (Dehaene-Lambertz et al., 2004), the tone onset time (Simos & Moltes, 1997), or the VOT (Kurtzberg, Stapells, & Wallace, 1988). These should not be neglected; the recording of a difference between physically different

standard and deviant stimuli is not sufficient to establish the existence of a mismatch response.³

Vigilance, attention toward the stimuli, number of repetitions of the standard before a deviant, size of the deviance, random or predictable presentation of the deviant, and the maturation stage of the network itself targeted by the experimental paradigm are other factors that differ across infant experiments. For instance, a salient or predictable deviant might induce more interactions between frontal selective attention mechanisms and auditory representations. Networks computing complex cues in which numerous assemblies of neurons are synchronized might imply larger cortico-cortical information flow within the network than within networks computing simple cues, especially during learning periods. Polarity of ERPs depending on the weight of activity in the different cortical layers and increase of activity in cortico-cortical connections might have an effect on the polarity of the response. In any case, mismatch negativity should not be considered as a flag that the brain exhibits when an auditory change is perceived and that infants possess or not, but mismatch responses should be understood as the result of complex interactions determined by the experimental conditions between input and output information flow within assemblies of neurons in a developing and learning brain. Further studies are needed in which all these factors should be systematically manipulated to understand the important contributing factors to record one or the other polarity over the frontal areas.

In summary, infants and adults exhibit similar behavioral performance in phoneme perception. ERPs reveal that this performance is subserved by similar computations of temporal neurons in both populations. Therefore, we hypothesize that in the case of phoneme processing, there is continuity between neonates and adults, and that from birth on infants are able to spontaneously compute phonemic representations. At the beginning of the life span, these phonemic representations are universal.⁴ This phonemic network, effective from the first days of life, is adequately configured to process the relevant properties of the speech environment and to detect any inherent regularities present in input. Exposed to a specific linguistic environment, the weights in the network are then modified to give rise to the adult phonemic representations without explicit reinforcement. A plausible explanation of this environmental effect is that infants exploit the statistical properties of language input (Maye et al., 2002; Kuhl, 2000; Holt, Lotto, & Kluender, 1998). Each language uses only a reduced repertoire out of the set of all possible phonemes, and productions of native speakers cluster around prototypes. For example, Lisker and Abramson (1964) reported that although variability exists within and between speakers, VOT values for the production of a particular phoneme tend to cluster around a mean value. If a phonemic feature is contrastive, tokens

present a bimodal distribution along this specific dimension, whereas if it is not contrastive, the distribution is unimodal (Maye et al., 2002). This particular distribution may distort the initial perceptive space, increasing the response to a change of category and decreasing the response to changes around the mean value. The result would be a magnet effect as suggested by Kuhl (2000), which would lead to a reduction of mismatch responses to foreign contrast as observed in adults and in infants as early as 1 year of age. In a recent article, Kuhl et al. (2003) showed that mere exposure to foreign speech is not in itself sufficient to modify the perceptual space. Social interactions are also necessary, either because social interactions direct infants' attention to the visual cues of speech production and thus consolidate phonemic representations through cross modal integration, or because a situation of social interaction may by itself activate linguistic networks, given that speech is the main communication medium in our species.

An argument often presented against a phonemic network is that animals show perceptual discontinuities similar to those observed in humans. For example, a mismatch response is recorded to a change in VOT from the brainstem of guinea pigs (King, McGee, Rubel, Nicol, & Kraus, 1995) and a common boundary along the VOT dimension is found in humans, macaques (Kuhl & Padden, 1983), and chinchillas (Kuhl & Miller, 1975). There is no reason to believe that if a strong acoustic discontinuity exists in perception, a linguistic network should ignore it. However, phonemic perception is not an exact mirror of acoustic perception. Using sine wave analogues of syllables, Serniclaes, Sprenger-Charolles, Carre, and Demonet (2001) explored discrimination functions along a /ba-/da/ continuum in children when the stimuli were perceived as electronic glissando and when they were perceived as speech. The location of the categorical boundary differed by one step between the speech mode and the nonspeech mode of perception. Along an /r-l/ continuum, humans and monkeys have different discrimination performance (Sinnott & Brown, 1997), whereas 2-month-old infants and English adults show the same categorical boundary (Eimas, 1975). On the other hand, because hearing is demonstrated during the last trimester of pregnancy, it is possible that the properties demonstrated in infants are the consequences of exposure to speech. However, the characteristics of the fetus environment would predict rather different properties for a learning network. The predominance of the mother voice inside the womb would favor precise representations of her productions and difficulties to normalize across different speakers. The perception of some phonemic contrasts, such as place of articulation, is very sensitive to noise. Recording within the uterus of a pregnant sheep, Griffiths, Brown, Gerhardt, Abrams, and Morris (1994) show that intelligibility was not good along this dimension. However, neonates have no problem discriminating phonemes such as /pa/ and /ta/

(Dehaene-Lambertz & Pena, 2001). Infants are also able to discriminate foreign contrasts not present in their environment (Best et al., 1988). These facts suggest that it is not exposure to speech that creates the capabilities described in infants even if the environment is certainly already shaping the phonemic representations.

In this article, we have emphasized the similarities between infants and adults, suggesting that phoneme perception relies on the same neural bases in infants and adults. However, we are far from understanding the precise mechanisms that govern the effect of the environment on this network. We do not know, for the present, how much exposure to a specific contrast is necessary, how long these environmental effects last without requiring new exposure to speech, whether visual cues are important, whether social interactions have the same impact in the first months of life and later on, or which exact statistical properties infants are sensitive to. Beyond phoneme perception, which we have discussed here, the recent progress in brain imaging techniques gives us access to the organization of the human brain from very early on and will permit us to better understand the ontogenesis of higher cognitive functions in the human brain.

Acknowledgments

This study was supported by grants from the Ministère Français de la Santé et de la Recherche PHRC 1995 No. AOM95011, Groupement d'Intérêt Scientifique Sciences de la Cognition No. PO 9004, ACI Blanche 1999 "Plasticité neuronale et acquisition du langage," and a McDonnell Foundation grant.

Reprint requests should be sent to Ghislaine Dehaene-Lambertz, Unité INSERM 562, Service Hospitalier Frédéric Joliot, CEA/DRM/DSV, 4 place du général Leclerc, 91401 Orsay cedex, France, or via e-mail: ghis@lscp.ehess.fr.

Notes

1. Linguists distinguish phones (which are physical categories) and phonemes (which are functional classes that distinguish words). Here, we use phoneme in a psychological sense, that is, as a perceptive category.
2. The MacGurk effect is related to the visual integration of incongruent articulatory movements during phoneme perception that induces a change of auditory perception although the auditory stimulus itself does not change. The ventriloquy effect is the attribution of the source of a sound to a person although the sound originates from a different location in space. Duplex perception is the integration of two stimuli into one single percept, although they are presented separately in each ear.
3. To avoid these problems, we have adapted the classical oddball paradigm. Stimuli are presented in trials of four stimuli each. Two types of trials, standard and deviant, are randomly presented. In standard trials, the four stimuli are identical (A A A A). In deviant trials, the last stimulus of the trial is different from the first three (B B B A). Across trials, A and B are randomly swapped, both stimuli being alternatively standard and deviant. Because the waveforms from the same syllable are compared, the observed difference can only be related to the context in which the last syllable was presented and to the

distance between this syllable and the context. This design has several important advantages over the classical oddball design. The introduction of the deviant stimulus is better controlled (it is always preceded by three standard stimuli). The response to the same stimulus at the same position in the trial is compared in standard and deviant trials, and the same number of trials is averaged in each condition. Because from trial to trial the context and test stimulus are exchanged, we can compare ERPs to the same stimulus when it is presented as standard and as deviant, and at a short time interval.

4. Universal in the sense that the same performance is shown by every human neonate and not in the sense that every phonemic boundary present across all languages of the world is already present at birth.

REFERENCES

- Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., & Näätänen, R. (1990). Event-related potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurology*, 77, 151–155.
- Alho, K., Sajaniemi, N., Niittyvuopio, T., Sainio, K., & Näätänen, R. (1990). ERPs to an auditory stimulus change in pre-term and full-term infants. *Psychophysiological Brain Research*, 2, 139–142.
- Altman, N. R., & Bernal, B. (2001). Brain activation in sedated children: Auditory and visual functional MR imaging. *Radiology*, 221, 56–63.
- Anderson, A. W., Marois, R., Colson, E. R., Peterson, B. S., Duncan, C. C., & Ehrenkranz, R. A. (2001). Neonatal auditory activation detected by functional magnetic resonance imaging. *Magnetic Resonance Imaging*, 19, 1–5.
- Barnet, A. B., Ohlrich, E. S., Weiss, I. P., & Shanks, B. (1975). Auditory evoked potentials during sleep in normal children from ten days to three years of age. *Electroencephalography and Clinical Neurophysiology*, 39, 29–41.
- Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology. General*, 117, 21–33.
- Bertrand, O., Perrin, F., & Pernier, J. (1985). A theoretical justification of the average reference in topographic evoked potential studies. *Electroencephalography and Clinical Neurophysiology*, 62, 462–464.
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 345–360.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., & Kaufman, J. N. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10, 512–528.
- Boatman, D., Lesser, R. P., & Gordon, B. (1995). Auditory speech processing in the left temporal lobe: An electrical interference study. *Brain and Language*, 51, 269–290.
- Bushnell, I. W. R. (1982). Discrimination of faces by young infants. *Journal of Experimental Child Psychology*, 33, 298–308.
- Caplan, D., Gow, D., & Makris, N. (1995). Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology*, 45, 293–298.
- Cheour, M., Alho, K., Ceponiene, R., Reinikainen, K., Sainio, K., Pohjavuori, M., Aaltonen, O., & Näätänen, R. (1998). Maturation of mismatch negativity in infants. *International Journal of Psychophysiology*, 29, 217–226.

- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, 1, 351–353.
- Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., Raz, A., Kaipio, M. L., Hiltunen, J., Aaltonen, O., Savela, J., & Hamalainen, H. (2002). Speech sounds learned by sleeping newborns. *Nature*, 415, 599–600.
- Colin, C., Radeau, M., Soquet, A., Demolin, D., Colin, F., & Deltenre, P. (2002). Mismatch negativity evoked by the McGurk–MacDonald effect: A phonetic representation within short-term memory. *Clinical Neurophysiology*, 113, 495–506.
- Courchesne, E. (1983). Cognitive components of the event-related brain potential: Changes associated with development. In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in event-related potential research: Endogenous components* (pp. 329–344). North-Holland: Amsterdam.
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, 8, 919–924.
- Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience*, 12, 449–460.
- Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain. *NeuroReport*, 9, 1885–1888.
- Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, 370, 292–295.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298, 2013–2015.
- Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience*, 12, 635–647.
- Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., & Dehaene, S. *Neural correlates of switching from auditory to speech perception*. Manuscript submitted for publication.
- Dehaene-Lambertz, G., & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *NeuroReport*, 12, 3155–3158.
- Dehaene-Lambertz, G., Pena, M., Christophe, A., & Landrieu, P. (2004). Phonetic processing in a neonate with a left sylvian infarct. *Brain and Language*, 88, 26–38.
- Dronkers, N. F., Redfern, B. B., & Knight, R. T. (2000). The neural architecture of language disorders. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 949–958). Cambridge: MIT Press.
- Duclaux, R., Challamel, M. J., Collet, L., Rouillet-Solignac, I., & Revol, M. (1991). Hemispheric asymmetry of late auditory evoked response induced by pitch changes in infants: Influence of sleep stages. *Brain Research*, 566, 152–158.
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1568–1578.
- Eggermont, J. J., & Ponton, C. W. (2002). The neurophysiology of auditory perception: From single units to evoked potentials. *Audiology & Neuro-otology*, 7, 71–99.
- Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the r–l distinction by young infants. *Perception and Psychophysics*, 18, 341–347.
- Eimas, P. D., & Miller, J. L. (1992). Organization in the perception of speech by young infants. *Psychological Science*, 3, 340–345.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.
- Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, 44, 33–66.
- Fowler, C. A., & Rosenblum, L. D. (1990). Duplex perception: A comparison of monosyllables and slamming doors. *Journal of Experimental Psychology. Human Perception and Performance*, 16, 742–754.
- Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *NeuroReport*, 13, 1251–1254.
- Gandour, J., Wong, D., Lowe, M., Dziedzic, M., Sathannuwong, N., & Tong, Y. (2002). A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience*, 14, 1076–1087.
- Giard, M. H., Lavikainen, J., Reinikainen, R., Perrin, F., Bertrand, O., & Pernier, J. (1995). Separate representations of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience*, 7, 133–143.
- Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds “L” and “R.” *Neuropsychologia*, 9, 317–323.
- Griffiths, S. K., Brown, W. S., Gerhardt, K. J., Abrams, R. M., & Morris, R. J. (1994). The perception of speech sounds recorded within the uterus of a pregnant sheep. *Journal of the Acoustical Society of America*, 96.
- Hauser, M. D., Dehaene, S., Dehaene-Lambertz, G., & Patalano, A. (2002). Spontaneous number discrimination of multi-format auditory stimuli in cotton-top tamarins (*Saguinus oedipus*). *Cognition*, 86, B23–B32.
- Holt, L. L., Lotto, A. J., & Kluender, K. R. (1998). Incorporating principles of general learning in theories of language acquisition. In M. Gruber, C. D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *Chicago linguistic society* (pp. 253–268). Chicago: Chicago Linguistic Society.
- Horvath, J., Czigler, I., Sussman, E., & Winkler, I. (2001). Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. *Brain Research, Cognitive Brain Research*, 12, 131–144.
- Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., & Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *Journal of Neuroscience*, 23, 9541–9546.
- Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630–645.
- Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, 43, 253–291.
- King, C., McGee, T., Rubel, E. W., Nicol, T., & Kraus, N. (1995). Acoustic features and acoustic change are represented by different central pathways. *Hearing Research*, 85, 45–52.
- Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage*, 13, 646–653.
- Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, 6, 263–285.

- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences, U.S.A.*, 97, 11850–11857.
- Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development*, 7, 361–381.
- Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced–voiceless distinction in alveolar plosive consonants. *Science*, 190, 69–72.
- Kuhl, P. K., & Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America*, 73, 1003–1010.
- Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, U.S.A.*, 100, 9096–9106.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science*, 255, 606–608.
- Kurtzberg, D., Stapells, D. R., & Wallace, I. F. (1988). Event-related potential assessment of auditory system integrity: Implications for language development. In P. M. Vietze & H. G. Vaughan (Eds.), *Early identification of infants with developmental disabilities* (pp. 160–179). Boston: Allyn and Bacon.
- Kurtzberg, D., Stone, C. L., & Vaughan, H. G. (1986). Cortical responses to speech sounds in the infant. In R. Cracco & I. Bodis-Wollner (Eds.), *Frontiers of Clinical Neuroscience, Vol. 3. Evoked potentials* (pp. 513–520). New York: Alan R. Liss.
- Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilaine, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *NeuroReport*, 13, 47–51.
- Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., & Näätänen, R. (2002). Maturation of the auditory change detection response in infants: A longitudinal ERP study. *NeuroReport*, 13, 1843–1848.
- Leppanen, P. H., Pihko, E., Eklund, K. M., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: II. Group effects. *NeuroReport*, 10, 969–973.
- Liberman, A. M., Isenberg, D., & Rakerd, B. (1981). Duplex perception of cues for stop consonants: Evidence for a phonetic mode. *Perception & Psychophysics*, 30, 133–143.
- Liebenthal, E., Binder, J. R., Piorkowski, R. L., & Remez, R. E. (2003). Short-term reorganization of auditory analysis induced by phonetic experience. *Journal of Cognitive Neuroscience*, 15, 549–558.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384–482.
- Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78, 91–121.
- Maurer, U., Bucher, K., Brem, S., & Brandeis, D. (2003). Development of the automatic mismatch response: From frontal positivity in kindergarten children to the mismatch negativity. *Clinical Neurophysiology*, 114, 808–817.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
- Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, 254, 1377–1379.
- Morr, M. L., Shafer, V. L., Kreuzer, J. A., & Kurtzberg, D. (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing*, 23, 118–136.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38, 1–21.
- Näätänen, R., Lehtokovski, A., Lennes, M., Cheour, M., Huotilainen, M., & Iivonen, A. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385, 432–434.
- Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, 11, 966–974.
- Nelson, C. A., & deRegnier, R. (1992). Neural correlates of attention and memory in the first year of life. *Developmental Neuropsychology*, 8, 119–134.
- Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan, J. H. G. (1989). Cortical responses to speech sounds and their formants in normal infants: Maturation sequences and spatiotemporal analysis. *Electroencephalography and Clinical Neurophysiology*, 73, 295–305.
- Pallier, C., Bosch, L., & Sebastian, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, 64, B9–B17.
- Parr, L. A., Winslow, J. T., Hopkins, W. D., & de Waal, F. B. (2000). Recognizing facial cues: Individual discrimination by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology*, 114, 47–60.
- Patterson, M. L., & Werker, J. F. (2003). Two-month old infants match phonetic information in lips and voice. *Developmental Science*, 6, 193–198.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., & Poeppel, D. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience*, 12, 1038–1055.
- Pihko, E., Leppanen, P. H., Eklund, K. M., Cheour, M., Guttorm, T. K., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: I. Age effects. *NeuroReport*, 10, 901–905.
- Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two components tones: Implications for voicing perception in stops. *Journal of the Acoustical Society of America*, 61, 1352–1361.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology. Human Perception and Performance*, 20, 421–435.
- Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., & Alku, P. (2001). Memory traces for words as revealed by the mismatch negativity. *NeuroImage*, 14, 607–616.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349–351.
- Remez, R. E., Pardo, J. S., Piorkowski, R. L., & Rubin, P. E. (2001). On the bistability of sine wave analogues of speech. *Psychological Science*, 12, 24–29.
- Rivera-Gaxiola, M., Csibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioural Brain Research*, 111, 13–23.
- Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997).

- The McGurk effect in infants. *Perception and Psychophysics*, 59, 347–357.
- Sams, M., Aulanko, R., Halalainen, M., Hari, M., Lounasmaa, O. V., & Lu, S. T. (1991). Seeing speech: Visual information from lip movement modified activity in the human auditory cortex. *Neuroscience Letters*, 127, 141–145.
- Sato, Y., Mori, K., Furuya, I., Hayashi, R., Minagawa-Kawai, Y., & Koizumi, T. (2003). Developmental changes in cerebral lateralization to spoken language in infants: Measured by near-infrared spectroscopy. *The Japan Journal of Logopedics and Phoniatrics*, 44, 165–171.
- Serniclaes, W., Sprenger-Charolles, L., Carre, R., & Demonet, J. F. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, 44, 384–399.
- Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *Journal of the Acoustical Society of America*, 106, 1078–1083.
- Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America*, 107(5 Pt 1), 2697–2703.
- Sharma, A., Marsh, C. M., & Dorman, M. F. (2000). Relationship between N1 evoked potential morphology and the perception of voicing. *Journal of the Acoustical Society of America*, 108, 3030–3035.
- Shtyrov, Y., & Pulvermüller, F. (2002). Memory traces for inflectional affixes as shown by mismatch negativity. *European Journal of Neuroscience*, 15, 1085–1091.
- Simos, P. G., & Moltes, D. L. (1997). Electrophysiological responses from a temporal order continuum in the newborn infant. *Neuropsychologia*, 35, 89–98.
- Sinnott, J. M., & Brown, C. H. (1997). Perception of the American English liquid /r-l/ contrast by humans and monkeys. *Journal of the Acoustical Society of America*, 102, 588–602.
- Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neuroscience Letters*, 266, 109–112.
- Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A magnetoencephalographic (MEG) study. *Neuroimage*, 9, 330–336.
- Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, 6, 391–398.
- Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience*, 13, 994–1005.
- Werker, J. F., & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganisation during the first year of life. *Infant Behavior and Development*, 7, 49–63.
- Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75, 1866–1878.
- Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349–355.
- Whalen, D. H., & Liberman, A. M. (1987). Speech perception takes precedence over nonspeech perception. *Science*, 237, 169–171.
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, 36, 638–642.
- Winkler, I., Kushnerenko, E., Horvath, J., Ceponiene, R., Fellman, V., & Huotilainen, M. (2003). Newborn infants can organize the auditory world. *Proceedings of the National Academy of Sciences, U.S.A.*, 100, 11812–11815.
- Woods, D. L., & Elmasian, R. (1986). The habituation of event-related potentials to speech sounds and tones. *Electroencephalography and Clinical Neurophysiology*, 66, 447–459.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6, 37–46.